Author's response to reviews

Title: The validity of using ICD-9 codes and pharmacy records to identify patients with chronic obstructive pulmonary disease

Authors:

Colin R Cooke (cookecr@umich.edu)
Min J Joo (joo@uic.edu)
Stephen M Anderson (stephen.anderson@va.gov)
Todd A Lee (todd.lee@va.gov)
Ed M Udris (Ed.Udris@va.gov)
Eric Johnson (johnson.ex@ghc.org)
David H Au (dau@u.washington.edu)

Version: 2 Date: 16 December 2010

Author's response to reviews: see over
To the Editorial Board:

Please accept the revised manuscript titled “The validity of using ICD-9 codes and pharmacy records to identify patients with chronic obstructive pulmonary disease” for re-consideration for publication as a Research Article in *BMC Health Services Research*. To the best of our ability, we addressed the concerns of the referees; the incorporation of their suggestions strengthened the manuscript. Below we present the itemized concerns of each referee in quoted italics. Our responses to each concern are then presented in plain text.

Several of the referees asked us to perform additional analyses that would require modifying the ICD-9 codes we used to identify COPD in the data source. The addition or removal of specific ICD-9 codes would certainly help readers understand the performance characteristics of each code which would be a valuable addition to the current manuscript. However, we regret that we no longer have the ability to parse the ICD-9 codes used in our administrative definition used in this analysis without re-pulling the data from its native form.

During the process of cohort development, programmers identified each hospital or outpatient encounter meeting our a priori defined criteria for a COPD-related visit (ICD-9 491.xx, 492.xx, 493.2, 496.xx). Unfortunately, the individual codes for each patient visit were not requested during the data pull. Thus we have no ability to examine each code or add or subtract codes in our data without re-initiating the data acquisition step again, which we cannot do. In the revised manuscript, we discuss this as a limitation.

Referee 1

1. “Age of study participants. I assume that as the data are collected from VA databases that all the participants are adults. And the authors have dichotomized age in the analysis as > or < 65 years old. However, was there a minimum age in the inclusion/exclusion criteria? COPD with airflow obstruction is really a disease of older individuals with demonstrated airflow obstruction. However, chronic bronchitis is a clinical phenotype of chronic productive cough with no mention of airflow limitation. As such young adults with chronic productive cough could be coded as “491” but not have COPD. Many researchers including ourselves have required a minimum age of 40 or 45 years to prevent including young adults with chronic cough.”
We did not exclude any patients based upon an age criterion because our aim was to be inclusive as possible and maintain generalizability. Although the literature is not clear about the lower limit of age where COPD is plausible, we agree that the specificity of ICD-9 codes for younger individuals is likely low. However, at the request of the reviewer we repeated our analyses excluding the all patients age < 40 years.

In total, 552 patients aged < 40 years old were excluded, 60 with airflow obstruction and 492 without airflow obstruction. To illustrate changes in model performance, we present the original areas under the ROC curve along with the areas under the ROC curve after exclusion of age < 40 in Table R1 below.

Table R1 – Model discrimination (area under the ROC curve) for original cohort and cohort after excluding patients < 40 years old.

<table>
<thead>
<tr>
<th></th>
<th>GOLD</th>
<th></th>
<th>LLN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC (95% CI)</td>
<td>Excluded age &lt; 40</td>
<td>AUC (95% CI)</td>
<td>Excluded age &lt; 40</td>
</tr>
<tr>
<td>1</td>
<td>≥ 1 outpatient ICD-9 code</td>
<td>0.75 (0.74-0.76)</td>
<td>0.74 (0.73-0.74)</td>
<td>0.71 (0.70-0.72)</td>
</tr>
<tr>
<td>2</td>
<td>≥ 2 outpatient ICD-9 code</td>
<td>0.75 (0.74-0.76)</td>
<td>0.74 (0.73-0.75)</td>
<td>0.72 (0.71-0.73)</td>
</tr>
<tr>
<td>3</td>
<td>≥ 3 outpatient ICD-9 codes</td>
<td>0.74 (0.73-0.75)</td>
<td>0.73 (0.72-0.74)</td>
<td>0.71 (0.70-0.72)</td>
</tr>
<tr>
<td>4</td>
<td>≥ 1 inpatient ICD-9 codes</td>
<td>0.62 (0.61-0.63)</td>
<td>0.60 (0.59-0.62)</td>
<td>0.57 (0.56-0.58)</td>
</tr>
<tr>
<td>5</td>
<td>≥ 1 outpatient ICD-9 codes + ≥ 1 Ipratropium Bromide MDI</td>
<td>0.77 (0.76-0.78)</td>
<td>0.76 (0.75-0.77)</td>
<td>0.75 (0.74-0.76)</td>
</tr>
<tr>
<td>6</td>
<td>≥ 1 outpatient ICD-9 codes + ≥ 1 albuterol MDI</td>
<td>0.76 (0.75-0.77)</td>
<td>0.75 (0.74-0.76)</td>
<td>0.74 (0.73-0.75)</td>
</tr>
<tr>
<td>7</td>
<td>≥ 1 outpatient ICD-9 code + ≥ 1 inpatient ICD-9 code + ≥ 3 Ipratropium bromide MDI + ≥ 6 albuterol MDI</td>
<td>0.78 (0.77-0.79)</td>
<td>0.77 (0.76-0.78)</td>
<td>0.76 (0.75-0.77)</td>
</tr>
<tr>
<td>8</td>
<td>Model 7 + age</td>
<td>0.79 (0.78-0.80)</td>
<td>0.77 (0.76-0.78)</td>
<td>0.77 (0.76-0.78)</td>
</tr>
<tr>
<td>9</td>
<td>Model 8 + smoking</td>
<td>0.79 (0.78-0.80)</td>
<td>0.78 (0.77-0.79)</td>
<td>0.77 (0.76-0.78)</td>
</tr>
</tbody>
</table>

By excluding patients < 40 years old there was no substantive changes in the primary results. In general, AUC slightly decreased due to patients less than 40 being much less likely to have airflow obstruction and less likely to have ICD-9 codes for COPD. Because results were similar, we elected to keep these patients in order to be inclusive. However, if the referee or editors feel strongly we would be willing to reconsider excluding these patients from the analysis.

Changes to manuscript: None.

2. “Many researchers have included ICD490 as one of the codes for COPD. However others, including our lab, have questioned the validity of this code for COPD. Did you consider ICD9 490 in your case definitions? Why or why not? How does this code change the sensitivity and specificity of the case definitions?”

We agree with the reviewer. We considered the inclusion of 490, but this code is non-specific for bronchitis and does not specify either acute or chronic. We ultimately opted to exclude this code in our administrative definition of COPD because of its questionable validity,[1] also noted by others [2]. As described above, we do not have the ability to test the performance of this code in our data, but describe our rationale for its exclusion in the revised manuscript.
Changes to manuscript: Page 5 in the Methods we state, “We did not include 490 in our administrative definition of COPD because the definition itself lacks specificity which increases the concern about misclassification.”

3. “Could the authors please define the difference between a ‘primary’ code and a ‘secondary’ one? I’m assuming that a secondary code occurs in additional coding fields available. In Canadian studies we often identify patients from the administrative hospital data if COPD occurs in any one of the 16 diagnostic fields, not just the first one. This is because for COPD exacerbations, the first code is often ‘influenza’ or ‘acute bronchitis’ while the second code is for COPD. We consider these patients to be COPD patients. Could the authors describe the primary and secondary coding issues and the implications of broadening the case definition to include secondary codes.”

The reviewer is correct. We clarified this in the manuscript text.

Changes to manuscript: Page 4 now states, “Outpatient primary and secondary ICD-9 codes were those recorded during a patient encounter in any outpatient clinic while inpatient primary ICD-9 codes were those recorded during an admission to the hospital. ICD-9 codes generated during visits to the pulmonary function laboratory were not considered in this analysis. Although secondary ICD-9 codes were considered for defining a COPD-related visit these were uncommonly coded by providers.”

4. “The authors selected the one prior to and one year after the index visit as the window for identifying COPD codes from the database. I’m a bit confused as to why the one year prior would provide strong results. Although it is possible that, for incident cases of COPD, a physician may provide a COPD code before the diagnosis has been confirmed with spirometry, in some cases the spirometry may rule out the diagnosis of COPD. In this situation, you have a COPD code for a visit prior to spirometry, then the spirometry does not show COPD so that would be a discordant case. I believe it would make more sense to have the year or two after the index/spirometry visit as the window for identifying COPD. Gershon and To used different year windows for their case definitions, such as 1 year, 5 years, or ever. What is the implication of using different time windows after the index spirometry?”

The reviewer makes an excellent point. Our hope was to provide a practical assessment of the validity of ICD-9 codes for COPD. Typically, administrative claims are scanned to identify patients with ICD-9 codes consistent with COPD without any ability to determine the temporal relationship between the date of encounter and the date of spirometry (if any was performed). Thus, our approach more approximates the way in which claims are used to identify cases. Although our decision to define the window around the spirometry date is arbitrary, we do not believe that this choice influences the validity of the study, particularly because empiric therapy COPD is not altered after spirometry is performed [3]. In addition, a unique feature of our study is that all patients completed spirometry. Without utilizing this event it would be difficult to define a cohort.

Changes to manuscript: None

5. “Although the authors mention it briefly in the limitations, there needs to be more discussion on how the very high prevalence seen in this study affected the positive predictive value. A prevalence of almost 50% is extremely high. I’m assuming this is due to the patient population – older men with substantial smoking exposure. Can the authors simulate what the PPV would be if the prevalence was more in line with population estimates of 10-20%?”

The reviewer correctly notes that the prevalence of COPD among in the VA, and our sample, is much greater than the general population [4]. This is also likely because our sample arose from patients who were referred for spirometry, rather than the general population. We added additional text to the
Discussion highlighting this point and added additional data presenting the PPV for the model presented in Table 3 when prevalence is either 10 or 20% in an online supplement. Because these data are simulated, we did not provide confidence intervals for these calculations.

**Changes to manuscript:** We included an additional table in the online supplement. Page 8, now states, “Because the prevalence of COPD in our cohort was high, we also estimated the positive and negative predictive values for the best performing model using prevalence estimates closer to that experienced in the general population (10-20%).”

Page 10 now reads, “Estimated PPV and NPV when the prevalence of COPD is 10 or 20% are presented in Additional file 1.”

6. “Was tiotropium available in this patient population, and if yes, how would the inclusion of this anticholinergic affect the results?”

Tiotropium was adopted quite slowly in the VA because of formulary restriction. At the time of our study it was rarely used.

**Changes to manuscript:** Page 6 now reads, “Tiotropium was not included in our analysis as it was adopted slowly in the VA because of formulary restriction.”

7. “In Table 1, could the authors include the GOLD stage breakdown instead of just mild/moderate/severe etc.”

**Changes to manuscript:** We added GOLD stage to Table 1.

8. “Table 1: BMI, for category one it isn’t 0 to 18.5 as your minimum cutoff was 15. Similarly, your 30+ category should be 30-55.”

**Changes to manuscript:** We corrected Table 1 to reflect the correct BMI categories.

9. “It seemed odd that the prevalence of congestive heart failure was the same in the two groups, as it is common comorbidity in COPD. Same comment with depression. Does the author have any comment on this – is it related to the VA population?”

Several studies among Veterans suggest that depressive disorders are more common among patients with COPD compared to those without. However, these studies relied upon clinical assessment for the diagnosis of depressive disorders. Although the specificity of ICD-9 based definitions of depressive disorders in the VA is excellent, the sensitivity is poor. Thus, the lower numbers of depression as defined by ICD-9 codes among patients with COPD may be a result of misclassification. [5]

**Changes to manuscript:** None.

10. “Figure 1 does not add much additional information and could be deleted. Perhaps a different figure regarding Item 1, next section below would be helpful.”

**Changes to manuscript:** We removed Figure 1 from the manuscript and placed it in an online supplement (additional file 1). We now detailed the cohort in the text of the manuscript.
11. “Some of the references are typed in inconsistently ie capital letters in all words in some titles but not others.”
12. “Reference 19 – physician is spelled incorrectly.”
13. “Reference 10 – Ontario should be capitalized. Page 5, second to last line: postbronchodilator spelled incorrectly”

Changes to manuscript: We corrected all references and the additional errors outlined above.

14. “It would be very informative to have a table with the frequency of the non-COPD codes that were used in the patients with COPD. If they were not being coded with COPD yet had airflow obstruction, what were they coded with? This is a question that is repeatedly raised – the authors have an opportunity to shed some light on this issue.”

The reviewer makes an excellent suggestion; however, as we describe above, we do not have the ability to evaluate the other ICD-9 codes for this cohort.

Changes to manuscript: No change.

Referee 2
1. “The authors use a large VA database to describe the validity of COPD-related ICD-9 codes when compared to objective test results (airway obstruction on post-BD spirometry). Since more than half of those with a COPD-related ICD-9 code had no airway obstruction, in the conclusion, change the phrase “likely misclassify a significant number...” (now understood to mean a statistical difference which is usually not clinically important) to something like “misclassifies the majority of patients...”

Changes to manuscript: The conclusion in the abstract and manuscript body now read, “Commonly used definitions of COPD in observational studies misclassify the majority of patients as having COPD.”

2. The use of two "gold standards" (a type of sensitivity analysis) is appropriate, since a large worldwide group of experts are calling for the industry-sponsored GOLD guidelines to discontinue their recommendation using post-BD FEV1/FVC <0.70 to define COPD since it causes very high COPD misclassification rates in older people. Two major ATS/ERS committees have disagreed for ten years regarding the definition of airway obstruction, so it should not be stated that using the fixed ratio to define COPD is "consistent with ATS/ERS standards.”

Changes to manuscript: We removed the phrase, “consistent with ATS/ERS standards”, and now state, “However, there continues to be disagreement among professional societies as to the optimal physiologic criteria to define COPD.”

3. I recommend labels of GOLD and LLN (instead of GOLD and Hankinson).

Changes to manuscript: We changed the “Hankinson” to “LLN” throughout the manuscript.

4. The gold standard LLN definition of COPD should require that both the post-BD FEV1/FVC and the FEV1 itself fall below their respective fifth percentile LLNs. A good argument could be made for using a definition of clinically-important COPD which includes only those with a post-BD FEV1 below 65% predicted (in the middle of GOLD stage two).”

The reviewer makes an excellent point. Although the clinical significance of each of the GOLD stages is currently under debate, our primary intent was to use the criterion standards (GOLD, LLN)
commonly accepted today regardless of their clinical relevance. This is how such codes are commonly used in observational studies. Thus, we felt that more specific definitions of COPD (e.g. clinically-important, GOLD II or greater) were out of the scope of this manuscript.

**Changes to manuscript:** None.

5. “Stratify model 1 into separate ICD-9 codes (or add separate models for each of the three primary COPD-related ICD-9 codes instead of only lumping them together). It is highly likely that the use of 491.xx chronic bronchitis has a much higher false positive rate for airway obstruction when compared to 496.xx chronic airway obstruction.”

We appreciate the reviewer’s concern about code 491.xx. However, as we describe above, we do not have the ability to determine the performance characteristics for each code.

**Changes to manuscript:** None.

6. “It is misleading to conclude that the models using GOLD definition performed better than the models using the LLN definition (which infers that the GOLD definition is more appropriate). The AUCs, Brier scores, and HL goodness of fit statistics were nearly identical and no statistical comparison between these models was done. If a stronger association between the fixed ratio and COPD-related ICD-9 codes actually exists, it is probably due to the inclusion (in both sides of the model) of patients with chronic bronchitis who have no airway obstruction (the O in COPD).

We appreciate the reviewers concern that our language could be interpreted as championing the GOLD criteria. This was not our intent. It is well known that GOLD runs the risk of misclassifying patients as having COPD when they do not [6]. Although the AUC is slightly better for the GOLD in our models, we appreciate that there is increasing support for the LLN standard as the preferred standard for the diagnosis of COPD.

**Changes to manuscript:** Our Discussions now states, “These variables showed similar performance when utilizing GOLD criteria for airflow obstruction compared to the LLN standard for airflow obstruction.”

7. Are the results for each of the 9 models in table 2 somehow combined for the pairs of models stratified by age group? Table 3 provides the results for model 8 for each age group separately.”

The reviewer highlights a point that needed clarification in the manuscript. Models presented in Table 2 include two variables for age. One is a continuous variable representing the patients age, the second is a dichotomous variable flagging each individual as >= 65 years old or < 65 years old. This latter dichotomous variable was interacted with all other variables in the model. This allowed separate coefficient estimates by age >=65 as well as single estimates of AUC, etc.

**Changes to manuscript:** We clarified model specification in the methods. Methods, page 7 now states, “We stratified all models by age ≥ 65 years by interacting age (≥65 years) with all variables in each model. This approach allowed separate coefficient estimates for patients ≥ 65 and <65 years of age allowing one to apply the model to Medicare and non-Medicare patients, but provides one number for each estimate of model performance (e.g AUC, Hosmer-Lemeshow, etc).”

8. “The coefficients for age would be easier to understand if given per decade (instead of per year).”
Changes to manuscript: We now present the coefficient for age per decade.

9. “Dose is spelled does in the abstract. References 7, 15, 18, and 37 are not complete (no page numbers). “

Changes to manuscript: Thank you for identifying these errors. We corrected these errors in the revised manuscript.

Referee 3

1. This is very interesting and well written manuscript. I have only a few minor comments that should, however, be considered before publication.

We appreciate the reviewer’s interest in our study.

Changes to manuscript: None.

2. Does the PFT test (that was performed for each patient in your data) "automatically" result in COPD inpatient or outpatient visit in your administrative data?

No. The encounter created by the interaction with the pulmonary function testing laboratory was not considered an encounter in our analysis.

Changes to manuscript: We clarified that the PFT encounter was not considered an encounter for the purposes of our study. Page 5 now reads, “Any ICD-9 codes generated during visits to the pulmonary function laboratory were not considered in this analysis.”

3. Exclusion of lung cancer patients should be feasible with administrative data, but how to do BMI-based exclusion with administrative data only? Should one try to take this into account while using administrative data to identify COPD?

The reviewer makes an excellent point that traditional administrative data does not include a measure of BMI. This exclusion only reduced the cohort by 68 patients. Inclusion of these individuals did not result in any substantive differences in our estimates. AUC for GOLD and LLN models were identical to the original estimates.

Changes to manuscript: We commented on the lack of change when BMI was not an exclusion criterion. Page 4 now state, “However, because BMI is not captured in most administrative data sources, we repeated all analyses including all patients regardless of BMI.” Page 10 now states, “Inclusion of all patients regardless of BMI resulted in no substantive changes in all models (data not shown).”

4. Was the data exactly +/- one year from index date, or +/- one “register years” (i.e. full calendar year before the index calendar year and full calendar year after the index year)? This is important, because it may affect the initial database queries to administrative registers (at least if applied in non-VA context.

We measured the one year pre- and one year post-index date as the 365 days occurring prior to spirometry and one year post spirometry. We clarified this in the methods.
Changes to manuscript: Page 5 now reads, “We collected demographic data, pharmacy records and the primary ICD-9 code for all outpatient and inpatient visits during the exact calendar date one year pre- and one year post the index date utilizing the VA computerized medical record system.

5. Justify more carefully why potentially COPD-related secondary diagnoses in outpatient visits were not considered as COPD visits.

Please see comment 2, reviewer 1 above. Additional outpatient codes beyond the primary codes were considered but are uncommon in the VA.

Changes to manuscript: Page 4 now states, “Outpatient primary and secondary ICD-9 codes were those recorded during a patient encounter in any outpatient clinic while inpatient primary ICD-9 codes were those recorded during an admission to the hospital. ICD-9 codes generated during visits to the pulmonary function laboratory were not considered in this analysis. Although secondary ICD-9 codes were considered for defining a COPD-related visit these were uncommonly coded by providers.”

6. Were comorbidities detected using only data prior to index date (and how many years "all previous outpatient visits" mean here)?

The reviewer is correct. Comorbidities were detected in the one year period prior to the index date. We clarified this point in the manuscript.

Changes to manuscript: Page 5 now reads, “Comorbid conditions relevant to patients with COPD were determined using ICD-9 codes for all previous outpatient visits in the one year period prior to the index date.”

7. Why only outpatient visit were used to detect comorbidities (and not inpatient admissions or pharmacy data)? Why didn’t you use (or refer to) widely applied techniques for detecting comorbidities such as Charlson or Elixhauser categories?

We chose to abstract specific comorbidities during data acquisition that are common among patients with COPD. We chose this approach to provide demographic data on the cohort to allow readers to assess the cohort’s generalizability. Scores such as Charlston or Elixhauser are traditionally used to provide a more comprehensive collection of comorbidities in the context of risk-adjustment. Pharmacy data may improve the sensitivity and specificity of the comorbidities, but this process requires a thorough and complex process of assigning medications to individual diseases. Because this approach was not consistent with our primary goal, we chose to rely upon specific ICD-9 codes alone for descriptive purposes. In addition, the minority of patients were hospitalized.

Changes to manuscript: None.

8. Please describe the structure of pharmacy data more precisely. Were those data describing prescriptions, purchases or reimbursements? What if there were more than one canister per prescription?

The Veterans Integrated Service Network (VISN) data warehouse contains the complete pharmacy records for patients who filled prescriptions within the VISN region. These data include the drug name, drug class, prescription identification number, prescription fill dates (for primary prescriptions and refills), number of allowable refills, date of next allowable refill, amount dispensed, day supply,
unit price of the medication and directions for use (the “sig”). This allowed us to precisely measure canister use and refills among all patients.

Changes to manuscript: We added the following text to page 5 of the methods, “The VISN data warehouse contains the complete pharmacy records for patients who filled prescriptions within the VISN region. These data include the drug name, class, prescription identification number, prescription fill dates (primary and refills), number of allowable refills, date of next allowable refill, amount dispensed, day supply, unit price of the medication and directions for use”.

9. In this study you had a fixed index date (PFT test). Do you think that the detection would work as well with an arbitrary index date for this population?

The reviewer makes an excellent point. Our decision to use the date of spirometry as the index date was arbitrary. We believe that the strategy for identifying a cohort would not make a substantive difference in our results, but we do not have the ability to test this hypothesis. Medication prescribing practice should not be impacted by our index date as is unlikely to differ pre and post spirometry (please see comment 5, Reviewer 1).

Changes to manuscript: None

10. Please discuss why you didn’t use (other) model selection criteria (such as AIC or MDL).

The primary goal for our analysis was to determine 1) the validity of ICD-9 codes for a physiologically defined diagnosis of COPD, and 2) the impact of pharmacy data on performance of a model including of ICD-9 codes. As such, we were not interested in optimizing performance and parsimony of our models during our analysis, or using a comprehensive list of variables to creating the best model. The AIC or MDL would be appropriate in selecting the model, between Models 7, 8 and 9, that has the best fit for a given complexity. We calculated AIC for these three models and Model 9 had the lowest AIC for both LLC and GOLD standard suggesting it is the best performing model.

Changes to manuscript: We reiterated that our intent was not to build the optimal model, but rather determine the impact of adding specific covariates to the performance of the model as measured by the AUC. Page 7 now reads, “Because model performance rather than parsimony was our primary concern, we did not employ measures such as the Akaike information criteria during the model building process.”

11. Please mention that AUC corresponds to C-statistics.

Changes to manuscript: We now describe that the AUC and C-statistic are synonymous.

12. Would it be useful to include information similar to Table 3 and 4 for other relevant models also as an online appendix?

We feel that presenting the best performing model is most relevant for the readers, but would be willing to include additional models in online only material if the reviewer or editors deems these as necessary.

Changes to manuscript: None.

Again, we believe that incorporating the reviewers’ suggestions has improved our manuscript. We look forward to hearing back from you.
Sincerely,

Colin R. Cooke, MD, MSc
Division of Pulmonary & Critical Care Medicine
Robert Wood Johnson Foundation Clinical Scholars Program
University of Michigan

References